



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Decoupling of DNA methylation and activity of intergenic LINE-1 promoters in colorectal cancer

Citation for published version:

Vafadar-Isfahani, N, Parr, C, McMillan, LE, Sanner, J, Yeo, Z, Saddington, S, Peacock, O, Cruickshanks, HA, Meehan, RR, Lund, JN & Tufarelli, C 2017, 'Decoupling of DNA methylation and activity of intergenic LINE-1 promoters in colorectal cancer', *Epigenetics*. <https://doi.org/10.1080/15592294.2017.1300729>

Digital Object Identifier (DOI):

[10.1080/15592294.2017.1300729](https://doi.org/10.1080/15592294.2017.1300729)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Epigenetics

Publisher Rights Statement:

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.

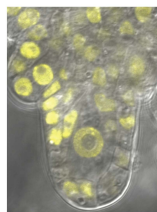
General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.





Decoupling of DNA methylation and activity of intergenic LINE-1 promoters in colorectal cancer

Natasha Vafadar-Isfahani, Christina Parr, Lara E. McMillan, Juliane Sanner, Zhao Yeo, Stephen Saddington, Oliver Peacock, Hazel A. Cruickshanks, Richard R. Meehan, Jonathan N. Lund & Cristina Tufarelli

To cite this article: Natasha Vafadar-Isfahani, Christina Parr, Lara E. McMillan, Juliane Sanner, Zhao Yeo, Stephen Saddington, Oliver Peacock, Hazel A. Cruickshanks, Richard R. Meehan, Jonathan N. Lund & Cristina Tufarelli (2017): Decoupling of DNA methylation and activity of intergenic LINE-1 promoters in colorectal cancer, Epigenetics, DOI: [10.1080/15592294.2017.1300729](https://doi.org/10.1080/15592294.2017.1300729)

To link to this article: <http://dx.doi.org/10.1080/15592294.2017.1300729>



View supplementary material



Accepted author version posted online: 16 Mar 2017.



Submit your article to this journal



Article views: 83



View related articles



View Crossmark data

Research Paper

Decoupling of DNA methylation and activity of intergenic LINE-1 promoters in colorectal cancer

Running title: Methylated and active LINE-1 promoters

Natasha Vafadar-Isfahani,¹ Christina Parr,¹ Lara E. McMillan,¹ Juliane Sanner,¹ Zhao Yeo,¹ Stephen Saddington,¹ Oliver Peacock,¹ Hazel A. Cruickshanks,² Richard R. Meehan,³ Jonathan N. Lund,^{1,*} and Cristina Tufarelli^{1,*}

¹School of Medicine, Royal Derby Hospital, University of Nottingham, Derby, UK, DE22 3DT

²School of Biological Sciences, University of Edinburgh, Edinburgh, UK, EH9 3FF

³MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK, EH4 2XU

*Corresponding author: Tel: +44 (0)1332 724704; Fax: +44 (0)1332 724626; Email: cristina.tufarelli@nottingham.ac.uk OR jon.lund@nottingham.ac.uk

Abstract

Hypomethylation of LINE-1 repeats in cancer has been proposed as the main mechanism behind their activation; this assumption, however, was based on findings from early studies that were biased towards young and transpositionally active elements. Here, we investigate the relationship between methylation of two intergenic, transpositionally inactive LINE-1 elements and

expression of the LINE-1 chimeric transcript (LCT) 13 and LCT14 driven by their antisense promoters (L1-ASP). Our data from DNA modification, expression, and 5'RACE analyses suggest that colorectal cancer methylation in the regions analyzed is not always associated with LCT repression. Consistent with this, in HCT116 colorectal cancer cells lacking DNA methyltransferases DNMT1 or DNMT3B, LCT13 expression decreases, while cells lacking both DNMTs or treated with the DNMT inhibitor 5-azacytidine (5-aza) show no change in LCT13 expression. Interestingly, levels of the H4K20me3 histone modification are inversely associated with LCT13 and LCT14 expression. Moreover, at these LINE-1s, H4K20me3 levels rather than DNA methylation seem to be good predictor of their sensitivity to 5-aza treatment. Therefore, by studying individual LINE-1 promoters we have shown that in some cases these promoters can be active without losing methylation; in addition, we provide evidence that other factors (e.g., H4K20me3 levels) play prominent roles in their regulation.

Key words

retrotransposons/ L1PA2/ LINE-1/ LCT/ colorectal cancer/ CRC/ L1-ASP/ methylation/ chromatin/ H4K20me3

List of abbreviations

LINE-1: long interspersed element 1

LCT: LINE-1 chimeric transcript

L1-ASP: LINE-1 antisense promoter

5-aza: 5-azacytidine

hMeDIP: hydroxymethylated DNA immunoprecipitation

5hmC: 5-hydroxymethylcytosine

ChIP: chromatin immunoprecipitation

Introduction

Long interspersed element 1 (LINE-1) regions are repetitive DNA sequences that comprise about 17% of the human genome, equivalent to approximately eight times the protein coding portion of the genome (1). LINE-1s are autonomous mobile elements that carry their own promoters and the information to copy and paste themselves to different locations in the genome, a phenomenon known as retrotransposition (2). Of the about 516,000 LINE-1s present in the human genome, just around 7000 have retained potentially intact promoters; of these, nearly 5000 are full-length but only up to 20 so called ‘hot-L1s’ have been shown to be able to retrotranspose (3-5).

It is now becoming more widely accepted that LINE-1s can also contribute to tumor in ways that are not related to their mobilization (6-8). Some of these effects (e.g., the ability of LINE-1

promoters to act as alternative promoters for protein coding genes (9)) are due to transcription from the LINE-1 promoters and therefore can be exerted by all elements carrying intact promoters and not restricted to the few hot-L1s. LINE-1s contain an internal bidirectional promoter within their 5' untranslated region (L1-5'UTR): a sense promoter (L1-SP) responsible for transcription of the element itself, and an antisense promoter (L1-ASP) driving transcription away from the element (10,11). L1-ASP activity has been mapped between positions 400 and 600 of the L1-5'UTR using transgenic constructs carrying retrotransposition competent LINE-1s (12). L1-ASP activity in this region was further confirmed in human embryonic stem and carcinoma cells (13). More recently, this activity has been shown to drive expression of the ORF0 protein with transposition enhancing properties and coded by about 781 LINE-1 loci in the human genome (14). Transcription initiation from L1-ASP in cancer has also been found between position 160 and 200 of the L1-5'UTR, in particular for intergenic, retrotransposition deficient LINE-1s (10,15).

It has been proposed that DNA methylation has evolved to prevent potentially damaging effects by silencing endogenous retrotransposons (16,17). Indeed, in somatic cells these elements are silenced and heavily methylated (18-20). By contrast, hypomethylation of LINE-1s is a common feature of cancer cells (21), which has been associated with LINE-1's activation and consequent further promotion of genome instability and cancer progression (22-24). In general, LINE-1 methylation levels decrease as the severity index of the cancer increases (25-27). LINE-1 activity and new somatic LINE-1 insertions have been described in a number of cancers of epithelial cell origin and happen at particularly high frequencies in colorectal tumors (28). These observations

have led to the notion that hypomethylation is the primary cause of activation of LINE-1 promoters in cancer. However, the evidence supporting this order of events is not conclusive.

Given their abundance and propensity to be methylated, LINE-1 methylation levels have been often used as a surrogate measure for global DNA methylation (29,30). However, until recently, the majority of LINE-1 methylation studies have looked at global levels using methods that favor detection of young and transpositionally active elements (31). These global studies lack resolution and assume one general regulatory network or machinery for all LINE-1 elements within the genome irrespective of cellular context. This is an unlikely scenario since LINE-1 elements can be found in a variety of loci and carry individual or familial SNPs. In agreement with this, there are numerous individual examples of LINE-1 elements that do not conform to the global methylation trends. Smith et al. (32) identified a subset of LINE-1 elements that were hypomethylated in oocytes despite the majority of transcription elements maintaining a heavy methylation status. In cancer, Phokaew et al. (33) showed that some LINE-1s are not affected by global hypomethylation. Similarly, generalized activation of LINE expression was not observed in mouse hypomethylation models (34,35).

These observations highlight the importance of studying individual LINE-1 elements in order to advance our understanding of their regulatory mechanism. Despite this, locus-specific LINE-1 studies are sparse and they are mainly focused on elements located within the body of annotated genes (intragenic LINE-1s) and their immediate effect on their host gene (36,37). We previously identified transcripts initiating at L1-ASP, which we referred to as LINE-1 chimeric transcripts or LCTs because they contain within the same molecule both LINE-1 and unique sequences (15).

Of particular interest, two LCTs (LCT13 and LCT14) were found to initiate at the antisense promoter (L1-ASP) of transpositionally inactive L1PA2 elements located at intergenic regions (between genes) on human chromosomes 7 and 5 respectively (15,38). We further demonstrated that in up to 50% of colorectal cancers lacking expression of the tumor suppressor gene *TFPI-2*, this silencing is associated with the presence of the overlapping anti-sense LCT13 transcript, suggesting that its expression may have a functional consequence in cancer (38). In an ES cell line model an antisense transcript can silence *TFPI-2* expression prior to de novo methylation of its promoter in differentiated cells (38). Our initial observations in MCF10A non-neoplastic breast cells suggested that in breast hypomethylation is involved in activation of the promoters of these intergenic L1PA2s (15); however, it remains unclear whether loss of DNA methylation is always necessary to activate them in cancer. In this study, we have investigated at a locus-specific level if loss of DNA methylation within the L1-5'UTR of the LINE-1s driving LCT13 and LCT14 is always associated with their expression, and we have also compared the presence of histone marks associated with active and inactive chromatin states to transcriptional activity and DNA methylation levels at these sites.

Results

Relationship between activity and methylation of LCT13 in cancer

We previously showed that aberrant activation of LINE-1 antisense promoters in cancer can occur at older and transposition deficient elements (15), such as the L1-ASP of the L1PA2 that drives expression of LCT13 (Figure 1A) (38). To determine whether this activity was associated with a decrease in methylation, we compared expression of LCT13 to its methylation in matched

normal and tumor tissues from six colorectal cancer (CRC) donors (26, 29, 30, 33, 65 and 104). In addition, three microsatellite instable (MSI: HCA-7, HCT116, RKO) and three microsatellite stable (MSS: CaCo-2; SW480 and SW620) CRC cell lines were tested as higher levels of methylation at LINE-1s have been reported in MSI relatively to MSS lines (36). The patients used for this study were chosen based on their LCT13 expression profile: 29 expresses LCT13 in the normal mucosa (N) but not in the tumor (T); 26 and 33 express LCT13 in N and at increased levels in T; 30, 65 and 104 expressed LCT13 only in T, the most commonly found pattern (Figure 1B, left panel) (38). Using the same quantitative RT-PCR approach in the CRC cell lines, we found that HCT116 express the highest levels of LCT13, followed by CaCo-2, HCA-7 and SW620; by contrast, LCT13 expression in RKO and SW480 is not detectable (Figure 1B, right panel). We then studied, by PCR on bisulfite treated DNA, the methylation of a region of the L1-5'UTR of the LINE-1 driving expression of LCT13 (bisulfite, Figure 1A middle diagram) in the same patient tissues and cell lines used for the expression analysis. With the exception of patients 33 and 30, where there was a decrease in methylation in the tumor tissue (T) samples, LCT13 is surprisingly highly methylated (>80%) in both normal (N) and T samples from all other patients studied (Figure 1C, left panel; Supplementary Figure 2A). Moreover, this L1-5'UTR region is also highly methylated in the six CRC cell lines, regardless of their expression profile (Figure 1C, right panel; Supplementary Figure 2A). Interestingly, in breast cancer (BC) cell lines the expected relationship between levels of methylation and expression of LCT13 was observed, with increasing amounts of expression corresponding to decreasing levels of methylation. However, whilst T47D are unmethylated, MCF7 and HCC1954 retain >55% methylation (Supplementary Figure 2B and C). These findings suggest that, at least in CRC, loss of

methylation may not be necessary for LINE-1 promoter activation, or, alternatively, that LCT13 expression in methylated tissues and cell lines may initiate at a different promoter.

LCT13 transcription is driven by L1-ASP

To address whether LCT13 transcription was driven by L1-ASP in the expressing cell lines, we performed 5'RACE anchored in the second exon of the LCT13 transcript (Figure 1A, bottom diagram). We first optimized conditions in MCF7 cells as we have previously been able to identify LCT13 transcription start sites (TSS) in these cells (15). The use of a variety of conditions highlighted the presence of weak and strong TSSs all within L1-5'UTR (Supplementary Figure 3). Using conditions optimal for the major TSS, we obtained 5'RACE products for all cell lines analyzed with the exception of RKO and SW480, consistent with the lack of expression of LCT13 in these two lines. In all positive cell lines LCT13 transcripts have a major transcription start site (TSS) within L1-5'UTR, confirming that L1-ASP drives LCT13 expression in these cells (Figure 2A). We next analyzed each TSS in the context of the average methylation of L1-5'UTR within the respective cell line. This revealed that the major TSS was situated within the methylated region of L1-5'UTR in HCT116 and MCF7, while in CaCo-2 and SW620 it lies just outside the CpG rich region studied by bisulfite sequencing (Figure 2B). In addition, HCC1954 has two TSSs that, judging by the relative amounts of 5'RACE products, seem to be used at similar levels: one TSS is in a similar location to that of MCF7 cells; the other is close to the TSS identified in HCT116 and T47D cells and situated within a methylated region (Figure 2B). This data suggest that, similar to many other RNA polymerase II promoters, the L1-ASP driving LCT13 has scattered TSSs and that in some cases (e.g., HCT116) transcription

initiation may occur despite the presence of DNA methylation; alternatively, the detection of methylation by bisulfite may indicate the presence of hydroxymethylated cytosine.

The co-occurrence of DNA methylation and expression of LCT13 and LCT14 is not due to increased levels of hydroxymethylcytosine

To address whether the observations made at LCT13 were unique to this particular locus, we analyzed the relationship between DNA methylation and expression of another intergenic LCT (LCT14), which we previously identified as initiating from the L1-ASP of an intergenic L1PA2 on chromosome 5 (Figure 3A) (15). We selected matched normal (N) and tumor (T) tissues from four CRC donors, two that did not show any LCT14 expression in N or T (148 and 218), one that expressed LCT14 in both N and T (153) and one that did not express LCT14 in N but did in T (205) (Figure 3B, left panel). DNA methylation levels were above 80% in all tissue samples, including those expressing LCT14 (Figure 3C left panel; Supplementary Figure 4A). We also compared LCT14 expression and methylation in five CRC and four BC cell lines. We could detect expression of LCT14 only in CaCo-2 and SW620 (Figure 3B, right panel); however, all CRC lines showed >85% methylation at this LCT (Figure 3C, right panel; Supplementary Figure 4A). In BC cells, we observed the predicted pattern of no expression and high methylation in HMEC, and high expression and no methylation in T47D; however, we saw expression of LCT14 in MCF7 cells that have levels of methylation comparable to those in HMEC, and some expression in HCC1954 with about 45% methylation (Supplementary Figure 4B and C). These observations are consistent with those made at LCT13 suggesting that at least in the context of CRC, transcription and methylation of LCT14 are not always mutually exclusive (Supplementary Figure 5). However, caution must be taken when interpreting these results because we used

bisulfite sequencing to assess the levels of methylation. It has been shown that this method is not able to distinguish between methylated (5mC) and hydroxymethylated (5hmC) cytosines, an intermediate product of cytosine demethylation (39). In order to address this issue, we performed hydroxymethylated DNA immunoprecipitations (hMeDIP) on the six CRC cell lines and on T47D cells that are not methylated at either LCTs. When compared to the levels detected for the hydroxymethylated DNA control, the levels of 5hmC at the two LCTs are very low in all cell lines (Figure 4), consistent with the reported loss in 5hmC in cultured cells (40). To better appreciate differences between the cell lines, we plotted the results removing the positive control and found that levels are very low, below 5% of Input, with slightly higher levels found at LCT14 than at LCT13 (Figure 4, inset). Interestingly, the three MSS lines and T47D have higher 5hmC levels when compared to the MSI lines, though these differences do not reach statistical significance for LCT13. It is important to note that T47D that has very low to undetectable DNA methylation by bisulfite sequencing at LCT14 shows levels of 5hmC that are comparable to those seen at SW480 which do not express LCT14 and have high levels of DNA methylation. Similarly, SW480 has the highest levels of 5hmC at LCT13, but do not express this LCT. These data indicate that there is no 5hmC enrichment at the L1-5'UTR of LCT13 and LCT14 suggesting that in some cell lines these LCTs can be active and methylated.

Effects of DNA methylation inhibition on LCT13 expression

To further investigate the relationship between methylation and expression of LCT13 we focused on HCT116 and RKO cells, established cell lines that have been widely used to study DNA methylation using inhibitors (41,42). We treated these cells with the DNA methylation inhibitor 5-azacytidine (5-aza) and confirmed that the treatment did not affect expression of the three

reference genes used for quantification of RT-PCR (Supplementary Figure 6A). Expression of *TFPI-2* upon 5-aza treatment, a gene known to be induced in cancer cells upon this treatment (43), confirmed its effectiveness (Supplementary Figure 6B). Comparison of LCT13 expression and methylation profiles in untreated, 5-aza or vehicle alone (DMSO) treated cells revealed that the 5-aza treatment has no effect on LCT13 expression (Figure 5A) or levels of DNA methylation (Figure 5B; Supplementary Figure 6C) in either cell lines.

The apparent lack of direct relationship between DNA methylation and regulation of LCT13 was further supported by the analysis of LCT13 expression in HCT116 lacking the maintenance (DNMT1) or *de novo* (DNMT3B) or both DNA methyltransferases (44,45). Surprisingly, a statistically significant decrease in the levels of LCT13 is seen in HCT116 cells lacking either DNMT1 or DNMT3B relatively to wild type cells whilst no significant differences are seen in HCT116 lacking both DNMTs (Figure 5C), consistent with the data from 5-aza treatment. Taken together these data suggest that in HCT116 cells regulation of the promoter of LCT13 relies on a more complex regulatory mechanism including factors other than DNA methylation.

Profile of histone modification at LCT13 and LCT14 in cancer cells

To gain a deeper insight into the epigenetic regulation of LCT13 and LCT14, we performed chromatin immunoprecipitation (ChIP) assays using antibodies raised against the active chromatin histone modification H3K4me3, and the heterochromatin associated marks H3K9me3, H3K27me3, and H4K20me3. We tested CRC cell lines HCT116 that express LCT13 but not LCT14, and RKO that are negative for both LCTs (Figures 1 and 3). As a comparison, we also analyzed the BC cell lines T47D and MCF-7, that express both LCTs but show no or some

methylation at the two L1-5'UTR, respectively (Supplementary Figures 2 and 4). The modification showing a more evident difference between expressed and non-expressed LCTs in the CRC lines is H4K20me3. LCTs that are not expressed (LCT14 in HCT116 and LCT13 & 14 in RKO) have an enrichment in H4K20me3 that is >25% Input (Figure 6, top and bottom left panels), a finding consistent with this modification being enriched at LINE-1 sequences (46). Surprisingly we found that enrichment in H3K4me3 at LCT13 in HCT116 cells was very modest, despite the fact that this LCT is expressed in these cells (Figure 1B); however, this finding is in agreement with the high levels of DNA methylation at LCT13 in HCT116 cells (Figure 1C) and with the reported mutual exclusivity of these two modifications (47,48). BC cells T47D that are not methylated at either loci, show the highest levels of H3K4me3 at both L1-5'UTRs (Figure 6, top right panel), while lower levels are seen in MCF-7 cells (Figure 6, bottom right panel). Levels of H4K20me3 at the LCT13 5'UTR are lower than those of H3K4me3 in both T47D and MCF-7 cells; by contrast, levels of H4K20me3 at the LCT14 5'UTR in both cell lines are comparable to those of H3K4me3 (Figure 6, top and bottom right panels). Interestingly, LCT13 that is not upregulated by 5-aza in RKO cells (Figure 4A) has low H3K4me3 and high H4K20me3 similar to what observed at the silenced LCT14 in HCT116 cells (Figure 6, top and bottom left panels). Instead, LCT14 that is expressed upon 5-aza treatment in RKO (Supplementary Figure 7A), displays comparable H3K4me3 and H4K20me3 levels of enrichment in these cells (Figure 6, bottom left panel), a pattern similar to that observed at the expressed LCT13 in HCT116 and at the expressed LCT14 in T47D and MCF-7 cells (Figure 6). Although no overall changes in methylation levels are observed between treated and untreated cells (Supplementary Figure 7B), LCT14 expression in 5-aza treated RKO cells is associated

with decrease in methylation at a specific CpG site (Supplementary Figure 7C). This data suggest that the relative enrichment in H4K20me3 at LCTs is better than DNA methylation at predicting whether a LINE-1 promoter is active or likely to be sensitive to 5-aza treatment.

Discussion

Cancer is a complex genetic disease, resulting from the accumulation of a number of genetic mutations and epigenetic changes that allow cells to overcome the normal biological hurdles and defense mechanisms limiting growth and division. Decreases in methylation at repetitive DNA is probably the most accepted paradigm among the global changes in cancer, so much so that loss of LINE-1 methylation is repeatedly used as a surrogate biomarker for measuring global DNA hypomethylation (49-52). Nevertheless, a significant decrease in LINE-1 methylation has not been seen in all cancer types. Susceptibility to LINE-1 hypomethylation appears to be cell and tissue specific and might be dependent on cancer type or subsets of a specific cancer type (53-55). Moreover, emerging evidence suggest that not all LINE-1s are regulated in the same way (56). Global LINE-1 methylation studies are unable to provide evidence for a relationship between methylation at individual LINE-1 and their activity. The observation that genes commonly downregulated in cancer are more likely to contain an intragenic LINE-1 indicates that the positioning of the LINE-1 is an important factor in its regulation (57). In agreement with genome-wide studies, loss of methylation at intronic LINE-1 situated within *c-MET*, *RAB3IP*, and *CHRM3* genes was found to lead to activation of transcription from L1-ASP (36,37,58). However, recent reports have shown that in mouse adult, embryonic and primordial germ cells loss of LINE-1 methylation does not always coincide with activation of transcription (34,35,59). In addition, the repressive role of DNA methylation at LINE-1 promoters was questioned in a

study in early development in mouse embryos that identified loss of the active histone H3K4me3 mark as the reason for decrease in L1 expression at the 8-cell stage compared to the 2-cell stage (60). Although these observations were made in mouse, which have LINE-1s that are different from humans, the majority of the studies in human colorectal cancer also point to the lack of a direct linear relationship between LINE-1 hypomethylation levels and tumor stage (27,61-63), suggesting that, in man and in mouse, DNA methylation cannot be the key regulator of expression from all LINE-1s.

In this study, methylation levels at the L1-5'UTR of two intergenic LINE-1s were investigated and compared to the levels of their specific LCT transcripts, LCT13 and LCT14. It must be noted that we were able to analyze region 1--436 of the L1-5'UTR of the two LCTs and, therefore, we did not include the L1-ASP promoter core activity that has been proposed to map between positions 450 and 600 of the 5'UTR (13,14). A direct correlation was not observed between LCT13 or LCT14 expression and methylation in the CRC patient samples, where, except for two individuals, LCT13 and LCT14 methylation levels were found to be high in both tumors and matched normal tissues regardless of their expression (Figures 1 and 3). The lack of a direct relationship between DNA methylation and promoter activity within the L1-5'UTRs of LCT13 and LCT14 was also observed in the breast and colon cancer cell lines analyzed, albeit in breast cells there was evidence of an inverse correlation between methylation and expression levels. Differential methylation and expression levels of LINE-1 elements in different cell lines and during different stages of development have been documented previously (64-66). However, it is also possible that changes in LINE-1 methylation levels can have different effects on an individual LINE-1 promoter activity depending not only on its unique location and surrounding

DNA but also on the cell type and whether the transcription machinery or silencing complexes are available. We have further characterized LCT13 by 5'RACE in these cell lines and we found that LCT13 transcription starts within L1-5'UTR in all expressing cell lines, and DNA methylation alone does not seem to be sufficient to prevent transcription (Figure 2). Consistent with a lack of relationship between methylation and expression of LCT13, 5-aza treatment did not appear to affect L1-ASP methylation levels nor LCT13 expression in HCT116 or RKO cells; however, caution must be taken when interpreting these data given that at the dose used there was no reduction of methylation at the LCT loci tested. Nevertheless, in agreement with the 5-aza findings, we found that HCT116 cells lacking both DNMT1 and DNMT3B show no changes in LCT13 expression relative to wild type cells; however, knock-out of either DNMTs led to a significant decrease in LCT13 expression (Figure 5).

Histone modification profiling at the LCT13 and LCT14 L1-5'UTRs (Figure 6) suggest H4K20me3 as a better indicator than DNA methylation to determine LCT expression or sensitivity to 5-aza. This modification has been previously shown to be enriched at LINE-1s in CRC (67), and a recent study has reported variation in the levels of H4K20me3 at different LINE-1 subtypes (68). We found that in HCT116 and RKO cells we could not induce expression of the LCTs that had a relative enrichment in H4K20me3 at levels above 40% of Input (LCT14 and LCT13, respectively) with 5-aza under the conditions used (Figure 5 and Supplementary Figure 7). Future experiments will be necessary to demonstrate the correlation between H4K20me3 and LCT activation, for example, by testing if loss of H4K20me3 (SUV4-20 knockdown) in cells in which LCTs are repressed is sufficient to upregulate expression. Moreover, extending the repertoire of unique LINE-1s analyzed will be necessary to confirm the

generality of this finding. This knowledge will be very important in the context of clinical applications of therapeutic agents such as 5-aza. For example, a therapy likely to promote activation of the LINE-1 promoter driving LCT13 or that intragenic to *cMET* could increase the risk of developing cancer by promoting silencing of the metastasis suppressor gene *TFPI-2* or expression of the oncogenic isoform L1-MET (37,38).

In conclusion, our data indicate that DNA methylation may not always play a key role in silencing the L1-5'UTRs of intergenic LINE-1s similar to those driving LCT13 and LCT14. It is more likely that the regulation of these promoters is achieved by a dynamic network and interplay between several factors that are compromised in some but not all cases of breast and colorectal tumors. Further investigations are needed to establish the other regulatory factors involved in LINE-1 promoter control and the exact role of DNA methylation in this process.

Materials and Methods

Patient details and ethics

Ethical approval for this study was granted by the Derbyshire Research Ethics Committee for the collection of normal and cancer colorectal tissues from patients who underwent surgical resection at the Royal Derby Hospital (Derby, UK) after informed written consent was obtained.

All tumors collected were large specimens of size ranging from 2.25 cm³ to 42 cm³; the sections used had $\geq 75\%$ tumor cells

Cell lines cultures and 5-azacytidine treatment

Cancer cell lines were obtained from ECACC (HCA-7, HCT116, CaCo2, SW620, SW480, MCF-7) or ATCC (RKO, HCC1954) and cultured according to the guidelines from the providers. HMEC (Invitrogen) were kindly provided by Dr. Allegrucci as frozen cell pellets ready for nucleic acid extraction. DNA samples from the cell lines were sent to Bio-Synthesis, Inc. for STR profiling to confirm cell line identity.

HCT116 and RKO Cells were treated for 72 hours with 1 μ M 5-azacytidine (Sigma, Cat. no. A2385) replacing the media every 24 hours with fresh media containing 1 μ M 5-aza. The HCT116 DNMT knockout cells were kindly donated by B. Vogelstein for the work published in (69) and their identity confirmed at that time via expression analysis and DNA methylation profiles. Cell pellets used for the analyses described in the present manuscript were prepared by Dr. Ottaviano from the same cultures, subsequently checked by expression analysis and DNA methylation profiles (manuscript in preparation).

Nucleic acids extractions and RT-PCR

DNA and total RNA was extracted from tissues and cell lines using Allprep RNA/DNA mini kit (QIAGEN, Cat. no. 80204) according to the manufacturer's instructions, including the on column DNaseI treatment (QIAGEN, Cat. no. 79254) for RNA extraction. RNA integrity was evaluated by agarose gel electrophoresis. Reverse transcription was carried out on 60 ng of RNA with random primers using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Cat. no. 4374966) in a final volume of 20 μ l and following manufacturer's instructions. An aliquot (2 μ l) of the reverse transcription reaction was used for real-time PCR

and each sample was analyzed in triplicate. Real time PCR was performed using custom designed Taqman assays for the LCTs and commercially available Taqman assays for the house keeping genes (Applied Biosystems). The thermocycling parameters were as follows: 10 minutes at 95°C, followed by 45 cycles of 15 seconds at 95°C, and 1 minute at 60°C. Details of all Taqman assays used can be found in Supplementary Table 1. Processing and analysis of real-time PCR data was carried out on GenEx software (bioMCC, Germany). Levels of LCT RNA expression were expressed relatively to the geometric mean of three reference genes (*HPRT*, *PGK*, and *GAPDH*; see Supplementary Methods and Supplementary Figure 1).

Bisulfite

Bisulfite conversion of DNA was carried out using EZ DNA Methylation-Gold Kit (Zymo research, Cat. no. D5005) according to the manufacturer's protocol. Primary and secondary PCRs were carried out using ZymoTaq (Zymo Research, Cat. no. E2001) in a final volume of 25 µl, under the following conditions: 10 min incubation at 95°C prior to 40 cycles of 30 sec at 95°C, 1 min at 48°C (49°C for TFPI2 primary primer pair) and 1 min at 72°C. Finally, the samples were incubated at 72°C for 10 min. Bisulfite treated DNA (40 ng) was used in the primary PCR reaction; 1 µl of the primary PCR was used as template for the secondary PCR. All primers were used at a final concentration of 0.6 µM (see Supplementary Table 2 for details of primers). PCR products were purified from the gel using QIAquick Gel Extraction kit (QIAGEN, Cat. no. 28704) and cloned using pGEM®-T Easy Vector System I (Promega, Cat. no. A1360). Transformation was carried out using XL10 Gold ultra-competent cells (Agilent Technologies, Cat. no. 200315) followed by selection on ampicillin (Sigma-Aldrich, Cat. no. A9518) containing plates following standard protocols. Plasmid purification was performed using

QIAprep Spin miniprep kit (QIAGEN, Cat. no. 27106) or PureYield™ Plasmid Miniprep System (Promega, Cat. no. A1223). Plasmid DNA was sent to Source Biosciences for Sanger sequencing. Sequencing data was analyzed using MacVector™ 8.0 and bisulfite analysis was carried out using the BIQ analyzer excluding clones with estimated conversion rates below 90% and clones 100% identical to another clone unless the overall methylation state was close to 0% or 100% (70).

5'RACE

5'RACE was carried out as previously described(15), except that the LCT13 reverse transcription primer (NV009: 5'TTTGTCCTTTTCTGTCAGGTCCTC3') and the gene specific primers for primary (NV010: 5'GCATCTTTTTGCCTGTTGTGGA3') and secondary (NV011: 5'ATCTTTTTGCCTGTTGTGGAGG3') PCR amplification were designed within exon 2 (Figure 1A, bottom diagram). *APRT* was used as a positive control using the previously published primers (15). PCR products were gel purified, cloned, and sequenced as described above.

hMeDIP

Immunoprecipitation of hydroxymethylated DNA was performed with a mouse monoclonal antibody against 5-hydroxymethylcytosine (5hmC) using the hMeDIP kit (Diagenode, Cat. no. C02010031) following the manufacturer's instructions. DNA samples were sheared at 4°C using a Diagenode Bioruptor. DNA (1 µg) was used for each IP; recovered DNA was analyzed by real-time PCR (Supplementary Table 3). Hydroxymethylated, methylated, and unmethylated DNA controls and primers for their amplification were provided with the kit.

Chromatin IPs

Chromatin Immunoprecipitation (ChIP) assays were performed using EZ-Magna ChIP™ A kit (Millipore, Cat. no. 17--408) according to the manufacturer's instructions. Antibodies raised in rabbit against H3K4Me3, H3K9me3, H3K27me3, and H4K20me3 were purchased from Diagenode (Cat. no. pAb-003-050, pAb-056-050, pAb-069-050 and pAb-057-050, respectively). Rabbit IgG was provided with the ChIP kit. Cells were crosslinked in their culture vessel with the appropriate media containing 1% formaldehyde (Sigma, Cat. no. F8775) for 10 minutes at room temperature. Cells were sonicated at 4°C using a Diagenode Bioruptor. Cells (10^6) were used for each IP and the recovered material was analyzed by real-time PCR using custom designed Taqman assays for LCT13, LCT14, and APRT (Supplementary Table 3). *GAPDH* primers were provided in the kit.

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

Conflict of interest

The authors declare no conflict of interest

Acknowledgements

We are grateful to Dr. Cinzia Allegrucci for donating human mammary epithelial cell pellets (HMEC) and Dr. Raffaele Ottaviano for providing the HCT116 single and double knockout cell pellets.

Financial Disclosure

Research was supported by the Derby Teaching Hospitals NHS Foundation Trust Colorectal Cancer Charitable Fund; and by Genetic Society summer scholarships to LMM and CP. NV-I was supported by a University Research Scholarship Doctoral Training Award from the University of Nottingham. Research in RRM lab is supported by the MRC. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Supplemental data for this article can be accessed on the publisher's website.

References

1. Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860-921.
2. Mills, R.E., Bennett, E.A., Iskow, R.C. and Devine, S.E. (2007) Which transposable elements are active in the human genome? *Trends Genet*, **23**, 183-191.
3. Beck, C.R., Collier, P., Macfarlane, C., Malig, M., Kidd, J.M., Eichler, E.E., Badge, R.M. and Moran, J.V. (2010) LINE-1 retrotransposition activity in human genomes. *Cell*, **141**, 1159-1170.
4. Brouha, B., Schustak, J., Badge, R.M., Lutz-Prigge, S., Farley, A.H., Moran, J.V. and Kazazian, H.H., Jr. (2003) Hot L1s account for the bulk of retrotransposition in the human population. *Proc Natl Acad Sci U S A*, **100**, 5280-5285.
5. Khan, H., Smit, A. and Boissinot, S. (2006) Molecular evolution and tempo of amplification of human LINE-1 retrotransposons since the origin of primates. *Genome Res*, **16**, 78-87.
6. Gifford, W.D., Pfaff, S.L. and Macfarlan, T.S. (2013) Transposable elements as genetic regulatory substrates in early development. *Trends in cell biology*, **23**, 218-226.
7. Goodier, J.L. and Kazazian, H.H., Jr. (2008) Retrotransposons revisited: the restraint and rehabilitation of parasites. *Cell*, **135**, 23-35.

8. Tufarelli, C. and Badge, R.M. (2016) In Cristofari, G. (ed.), *Human Retrotransposons in Health and Disease*. Springer, pp. 259-273.
9. Faulkner, G.J., Kimura, Y., Daub, C.O., Wani, S., Plessy, C., Irvine, K.M., Schroder, K., Cloonan, N., Steptoe, A.L., Lassmann, T. *et al.* (2009) The regulated retrotransposon transcriptome of mammalian cells. *Nat Genet*, **41**, 563-571.
10. Speek, M. (2001) Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. *Mol Cell Biol*, **21**, 1973-1985.
11. Swergold, G.D. (1990) Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol Cell Biol*, **10**, 6718-6729.
12. Yang, N. and Kazazian, H.H., Jr. (2006) L1 retrotransposition is suppressed by endogenously encoded small interfering RNAs in human cultured cells. *Nat Struct Mol Biol*, **13**, 763-771.
13. Macia, A., Munoz-Lopez, M., Cortes, J.L., Hastings, R.K., Morell, S., Lucena-Aguilar, G., Marchal, J.A., Badge, R.M. and Garcia-Perez, J.L. (2011) Epigenetic control of retrotransposon expression in human embryonic stem cells. *Mol Cell Biol*, **31**, 300-316.
14. Denli, A.M., Narvaiza, I., Kerman, B.E., Pena, M., Benner, C., Marchetto, M.C., Diedrich, J.K., Aslanian, A., Ma, J., Moresco, J.J. *et al.* (2015) Primate-specific ORF0 contributes to retrotransposon-mediated diversity. *Cell*, **163**, 583-593.

15. Cruickshanks, H.A. and Tufarelli, C. (2009) Isolation of cancer-specific chimeric transcripts induced by hypomethylation of the LINE-1 antisense promoter. *Genomics*, **94**, 397-406.
16. Jones, P.A. and Baylin, S.B. (2007) The epigenomics of cancer. *Cell*, **128**, 683-692.
17. Yoder, J.A., Walsh, C.P. and Bestor, T.H. (1997) Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet*, **13**, 335-340.
18. Harony, H., Bernes, S., Siman-Tov, R. and Ankri, S. (2006) DNA methylation and targeting of LINE retrotransposons in *Entamoeba histolytica* and *Entamoeba invadens*. *Mol Biochem Parasitol*, **147**, 55-63.
19. Hata, K. and Sakaki, Y. (1997) Identification of critical CpG sites for repression of L1 transcription by DNA methylation. *Gene*, **189**, 227-234.
20. Steinhoff, C. and Schulz, W.A. (2003) Transcriptional regulation of the human LINE-1 retrotransposon L1.2B. *Mol Genet Genomics*, **270**, 394-402.
21. Feinberg, A.P. and Vogelstein, B. (1983) Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature*, **301**, 89-92.
22. Eden, A., Gaudet, F., Waghmare, A. and Jaenisch, R. (2003) Chromosomal instability and tumors promoted by DNA hypomethylation. *Science*, **300**, 455.

23. Gaudet, F., Hodgson, J.G., Eden, A., Jackson-Grusby, L., Dausman, J., Gray, J.W., Leonhardt, H. and Jaenisch, R. (2003) Induction of tumors in mice by genomic hypomethylation. *Science*, **300**, 489-492.
24. Symer, D.E., Connelly, C., Szak, S.T., Caputo, E.M., Cost, G.J., Parmigiani, G. and Boeke, J.D. (2002) Human 11 retrotransposition is associated with genetic instability in vivo. *Cell*, **110**, 327-338.
25. Florl, A.R., Lower, R., Schmitz-Drager, B.J. and Schulz, W.A. (1999) DNA methylation and expression of LINE-1 and HERV-K provirus sequences in urothelial and renal cell carcinomas. *Br J Cancer*, **80**, 1312-1321.
26. Santourlidis, S., Florl, A., Ackermann, R., Wirtz, H.C. and Schulz, W.A. (1999) High frequency of alterations in DNA methylation in adenocarcinoma of the prostate. *Prostate*, **39**, 166-174.
27. Sunami, E., de Maat, M., Vu, A., Turner, R.R. and Hoon, D.S. (2011) LINE-1 hypomethylation during primary colon cancer progression. *PLoS One*, **6**, e18884.
28. Lee, E., Iskow, R., Yang, L., Gokcumen, O., Haseley, P., Luquette, L.J., 3rd, Lohr, J.G., Harris, C.C., Ding, L., Wilson, R.K. *et al.* (2012) Landscape of somatic retrotransposition in human cancers. *Science*, **337**, 967-971.
29. Kazazian, H.H., Jr. (2004) Mobile elements: drivers of genome evolution. *Science*, **303**, 1626-1632.

30. Deininger, P.L., Moran, J.V., Batzer, M.A. and Kazazian, H.H., Jr. (2003) Mobile elements and mammalian genome evolution. *Curr Opin Genet Dev*, **13**, 651-658.
31. Lisanti, S., Omar, W.A., Tomaszewski, B., De Prins, S., Jacobs, G., Koppen, G., Mathers, J.C. and Langie, S.A. (2013) Comparison of methods for quantification of global DNA methylation in human cells and tissues. *PLoS One*, **8**, e79044.
32. Smith, Z.D., Chan, M.M., Mikkelsen, T.S., Gu, H., Gnirke, A., Regev, A. and Meissner, A. (2012) A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature*, **484**, 339-344.
33. Phokaew, C., Kowudtitham, S., Subbalekha, K., Shuangshoti, S. and Mutirangura, A. (2008) LINE-1 methylation patterns of different loci in normal and cancerous cells. *Nucleic Acids Res*, **36**, 5704-5712.
34. Hackett, J.A., Reddington, J.P., Nestor, C.E., Dunican, D.S., Branco, M.R., Reichmann, J., Reik, W., Surani, M.A., Adams, I.R. and Meehan, R.R. (2012) Promoter DNA methylation couples genome-defence mechanisms to epigenetic reprogramming in the mouse germline. *Development*, **139**, 3623-3632.
35. Dunican, D.S., Cruickshanks, H.A., Suzuki, M., Semple, C.A., Davey, T., Arceci, R.J., Greally, J., Adams, I.R. and Meehan, R.R. (2013) Lsh regulates LTR retrotransposon repression independently of Dnmt3b function. *Genome biology*, **14**, R146.

36. Hur, K., Cejas, P., Feliu, J., Moreno-Rubio, J., Burgos, E., Boland, C.R. and Goel, A. (2013) Hypomethylation of long interspersed nuclear element-1 (LINE-1) leads to activation of proto-oncogenes in human colorectal cancer metastasis. *Gut*, **63**, 635-646.
37. Wolff, E.M., Byun, H.M., Han, H.F., Sharma, S., Nichols, P.W., Siegmund, K.D., Yang, A.S., Jones, P.A. and Liang, G. (2010) Hypomethylation of a LINE-1 promoter activates an alternate transcript of the MET oncogene in bladders with cancer. *PLoS Genet*, **6**, e1000917.
38. Cruickshanks, H.A., Vafadar-Isfahani, N., Dunican, D.S., Lee, A., Sproul, D., Lund, J.N., Meehan, R.R. and Tufarelli, C. (2013) Expression of a large LINE-1-driven antisense RNA is linked to epigenetic silencing of the metastasis suppressor gene TFPI-2 in cancer. *Nucleic Acids Res*, **41**, 6857-6869.
39. Nestor, C., Ruzov, A., Meehan, R. and Dunican, D. (2010) Enzymatic approaches and bisulfite sequencing cannot distinguish between 5-methylcytosine and 5-hydroxymethylcytosine in DNA. *Biotechniques*, **48**, 317-319.
40. Nestor, C.E., Ottaviano, R., Reddington, J., Sproul, D., Reinhardt, D., Dunican, D., Katz, E., Dixon, J.M., Harrison, D.J. and Meehan, R.R. (2012) Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Res*, **22**, 467-477.
41. Kim, J.G., Bae, J.H., Kim, J.A., Heo, K., Yang, K. and Yi, J.M. (2014) Combination effect of epigenetic regulation and ionizing radiation in colorectal cancer cells. *PLoS One*, **9**, e105405.

42. Weber, B., Kimhi, S., Howard, G., Eden, A. and Lyko, F. (2010) Demethylation of a LINE-1 antisense promoter in the cMet locus impairs Met signalling through induction of illegitimate transcription. *Oncogene*, **29**, 5775-5784.
43. Sato, N., Parker, A.R., Fukushima, N., Miyagi, Y., Iacobuzio-Donahue, C.A., Eshleman, J.R. and Goggins, M. (2005) Epigenetic inactivation of TFPI-2 as a common mechanism associated with growth and invasion of pancreatic ductal adenocarcinoma. *Oncogene*, **24**, 850-858.
44. Rhee, I., Jair, K.W., Yen, R.W., Lengauer, C., Herman, J.G., Kinzler, K.W., Vogelstein, B., Baylin, S.B. and Schuebel, K.E. (2000) CpG methylation is maintained in human cancer cells lacking DNMT1. *Nature*, **404**, 1003-1007.
45. Rhee, I., Bachman, K.E., Park, B.H., Jair, K.W., Yen, R.W., Schuebel, K.E., Cui, H., Feinberg, A.P., Lengauer, C., Kinzler, K.W. *et al.* (2002) DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature*, **416**, 552-556.
46. Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D. and Jenuwein, T. (2004) A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes & development*, **18**, 1251-1262.
47. Ooi, S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.P., Allis, C.D. *et al.* (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature*, **448**, 714-717.

48. Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Paabo, S., Rebhan, M. and Schubeler, D. (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet*, **39**, 457-466.
49. Link, A., Balaguer, F., Shen, Y., Lozano, J.J., Leung, H.C., Boland, C.R. and Goel, A. (2013) Curcumin modulates DNA methylation in colorectal cancer cells. *PLoS One*, **8**, e57709.
50. Schulz, W.A. (2006) L1 retrotransposons in human cancers. *J Biomed Biotechnol*, **2006**, 83672.
51. Tapp, H.S., Commane, D.M., Bradburn, D.M., Arasaradnam, R., Mathers, J.C., Johnson, I.T. and Belshaw, N.J. (2013) Nutritional factors and gender influence age-related DNA methylation in the human rectal mucosa. *Aging cell*, **12**, 148-155.
52. Wilson, A.S., Power, B.E. and Molloy, P.L. (2007) DNA hypomethylation and human diseases. *Biochimica et biophysica acta*, **1775**, 138-162.
53. Chalitchagorn, K., Shuangshoti, S., Hourpai, N., Kongruttanachok, N., Tangkijvanich, P., Thong-ngam, D., Voravud, N., Sriuranpong, V. and Mutirangura, A. (2004) Distinctive pattern of LINE-1 methylation level in normal tissues and the association with carcinogenesis. *Oncogene*, **23**, 8841-8846.
54. Florl, A.R., Steinhoff, C., Muller, M., Seifert, H.H., Hader, C., Engers, R., Ackermann, R. and Schulz, W.A. (2004) Coordinate hypermethylation at specific genes in prostate carcinoma precedes LINE-1 hypomethylation. *Br J Cancer*, **91**, 985-994.

55. Thayer, R.E., Singer, M.F. and Fanning, T.G. (1993) Undermethylation of specific LINE-1 sequences in human cells producing a LINE-1-encoded protein. *Gene*, **133**, 273-277.
56. Philippe, C., Vargas-Landin, D.B., Doucet, A.J., van Essen, D., Vera-Otarola, J., Kuciak, M., Corbin, A., Nigumann, P. and Cristofari, G. (2016) Activation of individual L1 retrotransposon instances is restricted to cell-type dependent permissive loci. *eLife*, **5**, e13926.
57. Apornthewan, C., Phokaew, C., Piriyaongsa, J., Ngamphiw, C., Ittiwut, C., Tongsimas, S. and Mutirangura, A. (2011) Hypomethylation of intragenic LINE-1 represses transcription in cancer cells through AGO2. *PLoS One*, **6**, e17934.
58. Roman-Gomez, J., Jimenez-Velasco, A., Agirre, X., Cervantes, F., Sanchez, J., Garate, L., Barrios, M., Castillejo, J.A., Navarro, G., Colomer, D. *et al.* (2005) Promoter hypomethylation of the LINE-1 retrotransposable elements activates sense/antisense transcription and marks the progression of chronic myeloid leukemia. *Oncogene*, **24**, 7213-7223.
59. Seisenberger, S., Andrews, S., Krueger, F., Arand, J., Walter, J., Santos, F., Popp, C., Thienpont, B., Dean, W. and Reik, W. (2012) The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Molecular cell*, **48**, 849-862.
60. Fadloun, A., Le Gras, S., Jost, B., Ziegler-Birling, C., Takahashi, H., Gorab, E., Carninci, P. and Torres-Padilla, M.E. (2013) Chromatin signatures and retrotransposon profiling in mouse embryos reveal regulation of LINE-1 by RNA. *Nat Struct Mol Biol*, **20**, 332-338.

61. Matsunoki, A., Kawakami, K., Kotake, M., Kaneko, M., Kitamura, H., Ooi, A., Watanabe, G. and Minamoto, T. (2012) LINE-1 methylation shows little intra-patient heterogeneity in primary and synchronous metastatic colorectal cancer. *BMC cancer*, **12**, 574.
62. Murata, A., Baba, Y., Watanabe, M., Shigaki, H., Miyake, K., Ishimoto, T., Iwatsuki, M., Iwagami, S., Sakamoto, Y., Miyamoto, Y. *et al.* (2013) Methylation levels of LINE-1 in primary lesion and matched metastatic lesions of colorectal cancer. *British journal of cancer*, **109**, 408-415.
63. Ogino, S., Nosho, K., Kirkner, G.J., Kawasaki, T., Chan, A.T., Schernhammer, E.S., Giovannucci, E.L. and Fuchs, C.S. (2008) A cohort study of tumoral LINE-1 hypomethylation and prognosis in colon cancer. *J Natl Cancer Inst*, **100**, 1734-1738.
64. Belancio, V.P., Hedges, D.J. and Deininger, P. (2006) LINE-1 RNA splicing and influences on mammalian gene expression. *Nucleic Acids Res*, **34**, 1512-1521.
65. Ergun, S., Buschmann, C., Heukeshoven, J., Dammann, K., Schnieders, F., Lauke, H., Chalajour, F., Kilic, N., Stratling, W.H. and Schumann, G.G. (2004) Cell type-specific expression of LINE-1 open reading frames 1 and 2 in fetal and adult human tissues. *J Biol Chem*, **279**, 27753-27763.
66. Kines, K.J. and Belancio, V.P. (2012) Expressing genes do not forget their LINES: transposable elements and gene expression. *Frontiers in bioscience*, **17**, 1329-1344.

67. Gezer, U., Ustek, D., Yoruker, E.E., Cakiris, A., Abaci, N., Leszinski, G., Dalay, N. and Holdenrieder, S. (2013) Characterization of H3K9me3- and H4K20me3-associated circulating nucleosomal DNA by high-throughput sequencing in colorectal cancer. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*, **34**, 329-336.
68. Nelson, D.M., Jaber-Hijazi, F., Cole, J.J., Robertson, N.A., Pawlikowski, J.S., Norris, K.T., Criscione, S.W., Pchelintsev, N.A., Piscitello, D., Stong, N. *et al.* (2016) Mapping H4K20me3 onto the chromatin landscape of senescent cells indicates a function in control of cell senescence and tumor suppression through preservation of genetic and epigenetic stability. *Genome biology*, **17**, 158.
69. Sproul, D., Nestor, C., Culley, J., Dickson, J.H., Dixon, J.M., Harrison, D.J., Meehan, R.R., Sims, A.H. and Ramsahoye, B.H. (2011) Transcriptionally repressed genes become aberrantly methylated and distinguish tumors of different lineages in breast cancer. *Proc Natl Acad Sci U S A*, **108**, 4364-4369.
70. Bock, C., Reither, S., Mikeska, T., Paulsen, M., Walter, J. and Lengauer, T. (2005) BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. *Bioinformatics*, **21**, 4067-4068.

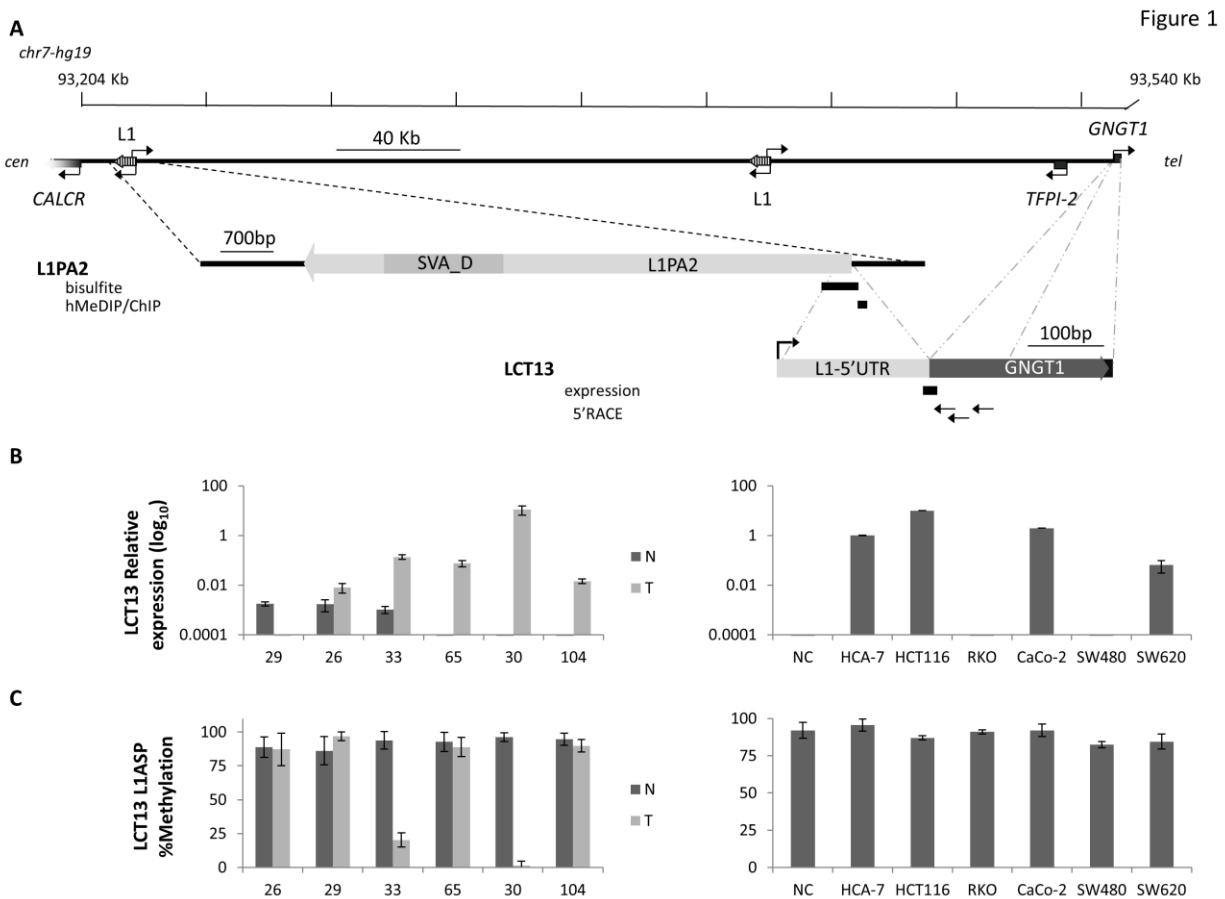


Figure 1. Relationship between methylation and expression of LCT13 L1ASP in CRC. (A)

Top: Schematic diagram of the LCT13 genomic locus on human chromosome 7 (chr7:93,204,042-93,540,485; center) with indicated the positions of the *CALCR*, *TFPI-2*, and *GNGT1* genes and of the two intact intergenic LINE1s (L1) present in this region. Middle: enlargement of the LINE-1 (L1PA2: chr7:93,213,393-93,221,079, with an SVA_D spanning the interval 93,214,544-93,216,214) from which LCT13 originates with the regions (black bars) tested by bisulfite or hMeDIP and ChIP. Bottom: enlargement of the LCT13 spliced transcript with indicated its exon structure [LINE-1 5'UTR fragment in light grey (chr7:93,220,882-93,221,083) and, in dark grey, the two GNGT1 exons (93,536,051-93,536,154 and 93,540,102-

93,540,235), part of the LCT13 transcript]. Also indicated are the positions of the Taqman assay used for LCT13 expression studies located at the splice junction (black bar) and of the primers used for 5'RACE (arrows). All coordinates are from hg19 annotations; scale is in kilobase pairs (kb). **(B)** Bar charts showing the expression of LCT13 measured by real time RT-PCR and expressed relatively to the geometric mean of three reference genes in matched normal (dark grey, N) and tumor (light grey, T) tissues from six colorectal cancer patients (left panel) and six cell lines (right panel). NC: normal colon, commercially sourced total RNA from seven healthy donors pooled together. **(C)** Bar charts of the methylation levels measured by bisulfite sequencing in the tissues of the six patients and cell lines presented in B.

Figure 2

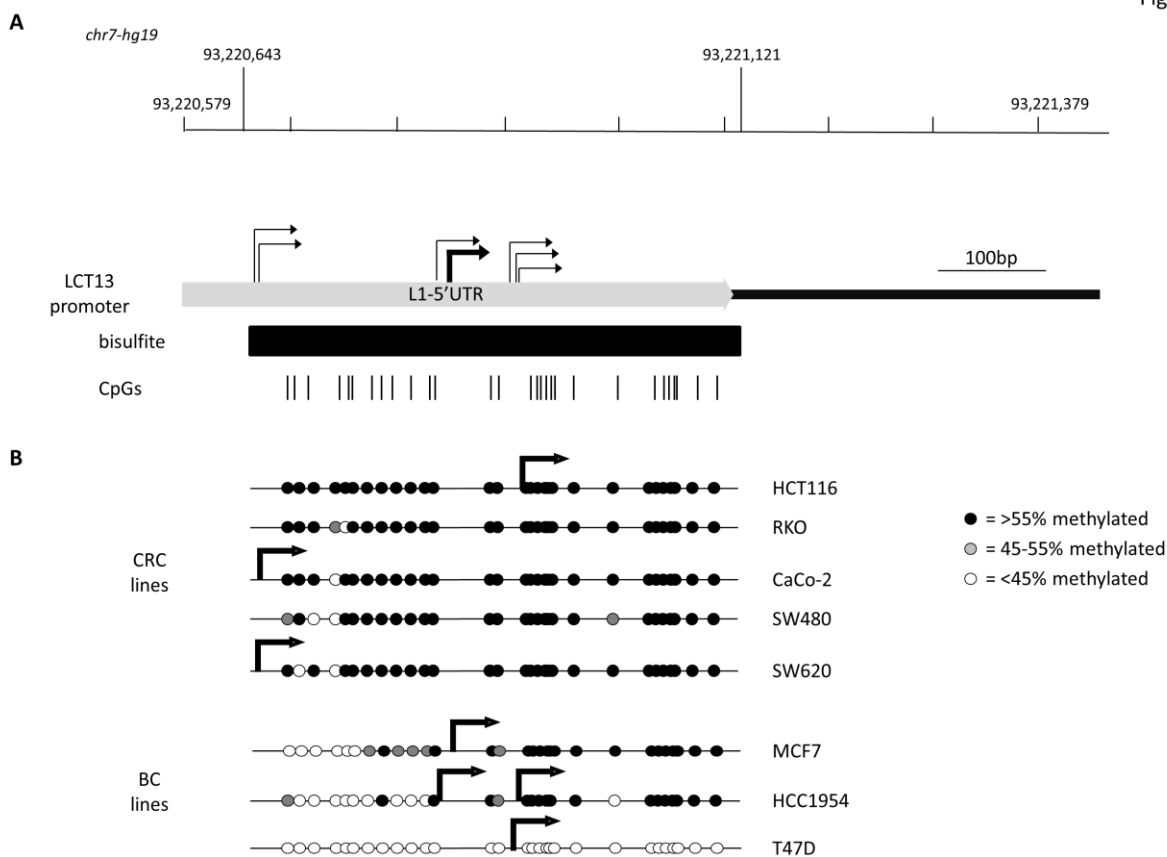


Figure 2. Relationship between L1-5'UTR methylation and LCT13 transcription start sites in cell lines. (A) Schematic diagram showing the 5'UTR of the L1PA2 driving LCT13 (chr7: 93,220,579-93,221,079) with indicated the regions analyzed by bisulfite sequencing (black bar; chr7: 93,220,643-93,221,121) and the positions of the 29 CpG sites (vertical black lines) within it. All coordinates are from hg19 annotations and the scale is in base pairs (bp). Indicated are all the transcription start sites (TSS; bent arrows) identified in the cell lines by 5'RACE demonstrating scattered transcription initiation (light grey) (see also Supplementary Figure 3). (B) Diagrams combining the lollipops summarizing the average methylation at each of the 29 CpG site analyzed in the panels of colorectal (CRC) and breast (BC) cancer cell lines (see

Supplementary Figure 2) with the stronger TSS site identified for the particular cell line (thick bent arrows). No TSS was identified by 5'RACE in RKO and SW480 cell lines, consistent with lack of detectable LCT13 transcripts in these cells (Figure 1B, right panel).

Accepted Manuscript

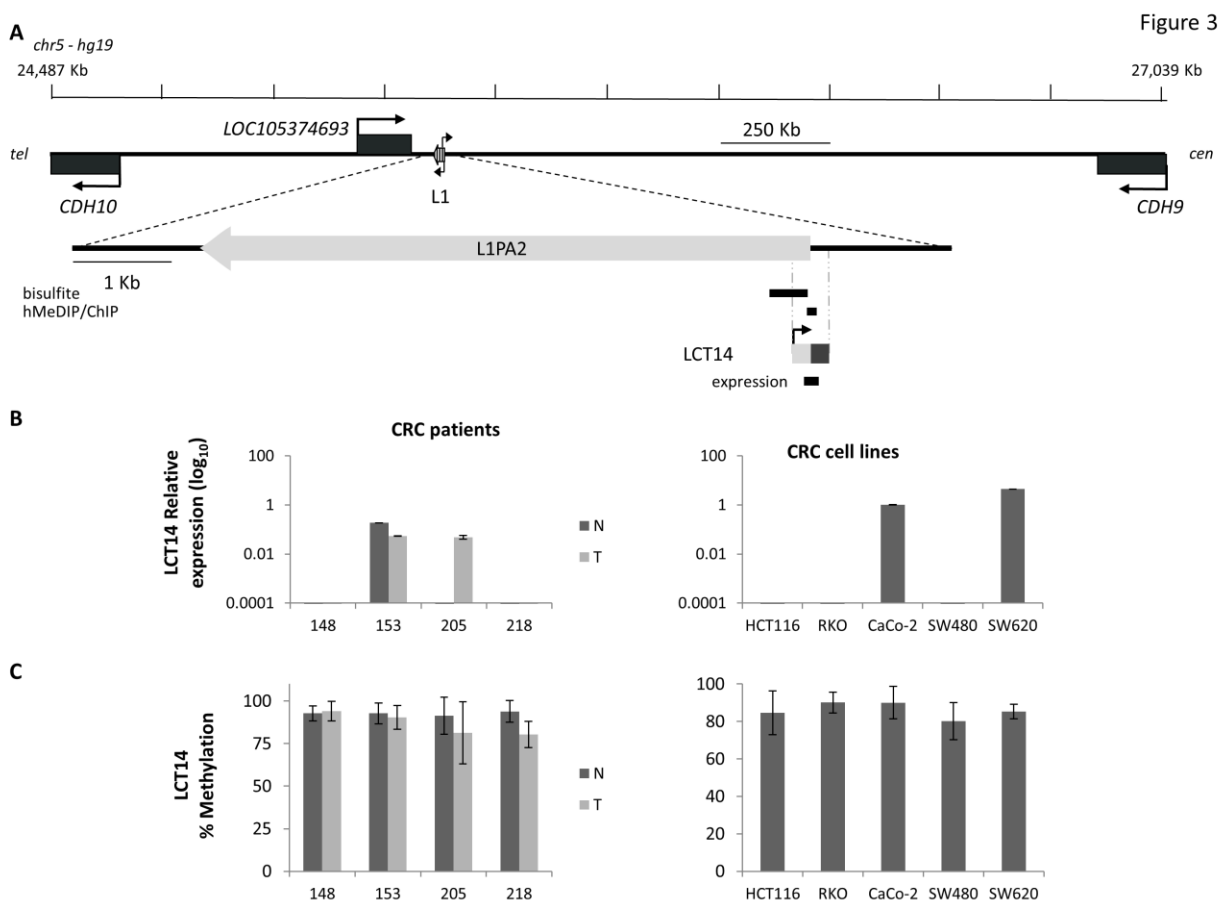


Figure 3. Relationship between methylation and expression of LCT14 in CRC. (A)

Schematic diagram of the LCT14 genomic locus on human chromosome 5 (coordinates: 24,487,209-27,038,689) with indicated positions of the annotated genes *CDH10*, *LOC105374693*, and *CDH9* and of the intact intergenic LINE1 (L1) that drives transcription of LCT14. At the bottom is an enlargement of the region including the LINE-1 (L1PA2; chr5:25,378,639-25,384,665) from which LCT14 originates with the regions (black bars) tested by bisulfite or hydroxymethylated DNA (hMeDIP) and chromatin (ChIP) immunoprecipitations and, below these, the LCT14 transcript (chr5: 25,384,485-25,384,958) and the region amplified for expression studies. All coordinates are from hg19 annotations; scale is in kb. **(B)** Expression

of LCT14 measured by real-time RT-PCR and expressed relatively to the geometric mean of three reference genes in matched normal (dark grey) and tumor (light grey) tissues from four colorectal cancer patients (left panel) and of five colorectal cancer cell lines (right panel). (C) Methylation levels measured by bisulfite sequencing in the paired normal and tumor tissues of the four patients (left panel) and cell lines (right panel) described in B.

Figure 4

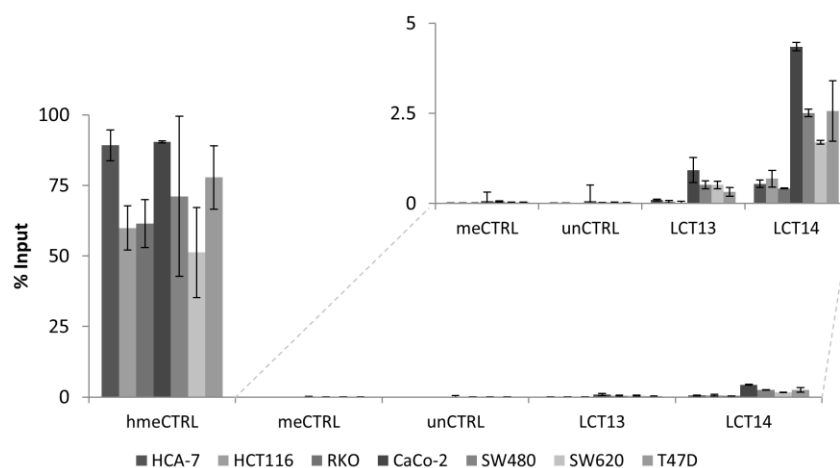
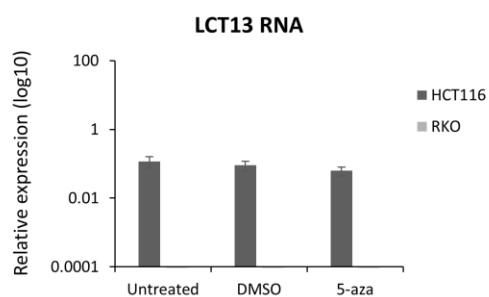
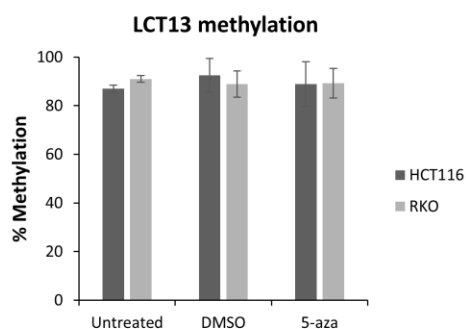


Figure 4. Analysis of 5hmC at LCT13 and LCT14 in cancer cell lines. Levels of hydroxymethylcytosine (hmeC) obtained by hMeDIP and expressed as % of Input. hmeCTRL: hydroxymethylated control DNA; meCTRL: methylated control DNA; unCTRL: unmethylated control DNA. The inset shows an enlargement of the region of the graph without the positive hmeCTRL indicating that some cell lines (Caco-2, SW480, SW620 and T47D) show minor enrichment at LCT13 and LCT14 L1ASPs, relative to the negative controls and the negative cell lines.

A



B



C

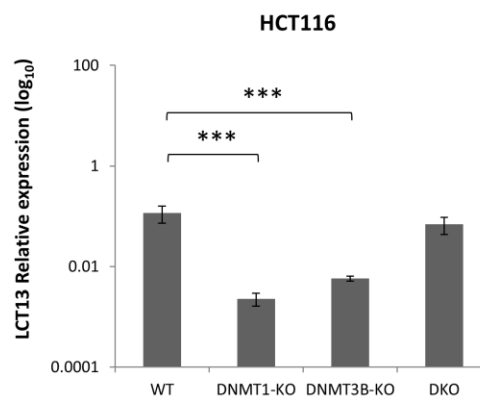


Figure 5. Effects of DNA methylation inhibition in colon cancer cell lines. Expression (A) and methylation (B) of LCT13 in HCT116 cells untreated or treated with DMSO vehicle alone or 1 μ M 5-aza in DMSO. Treatment with 5-aza has no effect on LCT13 expression levels; no overall changes in the levels of DNA methylation are seen in either cell line. (C) Expression profile of LCT13 in HCT116 that are either wild type or lacking DNA methyltransferase 1 (DNMT1-KO), or 3B (DNMT3B-KO), or both DNMT1 and 3B (DKO). *P* values were calculated by one-way ANOVA. ***: *P*<0.001

Figure 6

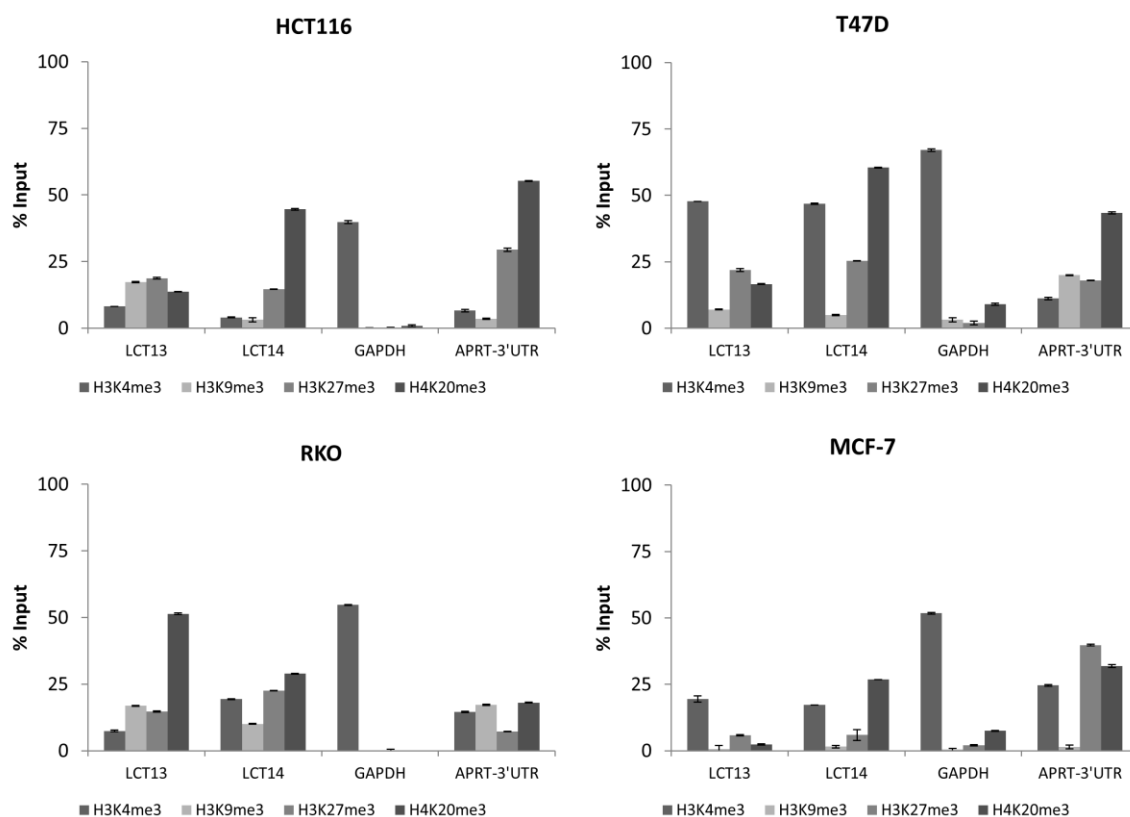


Figure 6. Histone modifications at LCT13 and LCT14 in cancer cell lines. ChIP assays performed using antibodies against active (H3K4me3) and repressive (H3K9me3, H3K27me3, H4K20me3) histone marks in LCT13 positive and LCT14 negative HCT116 cells (top left panel), in LCT13 and LCT14 negative RKO cells (bottom left panel), and in LCT13 and LCT14 positive T47D (top right panel) and MCF-7 (bottom right panel) cells. GAPDH is a promoter Taqman assay used as a positive control for H3K4me3 and APRT-3'UTR has been previously shown to be enriched at repressive mark H4K20me3.